DHRUV KUMAR

CONTACT INFORMATION

4-192 Keller Hall, 200 Union St. SE
Minneapolis, MN
United States - 55414

Mobile: (1) 651-955-8923
E-mail: kumar434@umn.edu
URL: **www.dhruvkumar.me**

RESEARCH INTERESTS

Distributed Systems, Parallel Computing, Data Mining, Machine Learning.

EDUCATION

University of Minnesota, Twin Cities, United States **Sep 2017 - Present**
PhD in Computer Science

- CGPA: **4.0 / 4.0**
- 3M Science and Technology Fellowship.
- **Relevant Courses**: Distributed Systems, Advanced Machine Learning, Matrix Theory, Probability and Statistics.

Birla Institute of Technology and Science (BITS), Pilani, India **Aug 2010 - May 2014**
Bachelor of Engineering (Hons.) Computer Science

- CGPA: **9.92 / 10.0**
- **Rank 1** in Class of 2014 of Computer Science, comprising of 120 students.
- **Rank 3** in Class of 2014 of BITS-Pilani, comprising of 800 students.
- **Relevant Courses**: Data Structures, Algorithms, Computer Organization, Operating Systems, Database Systems, Computer Networks, Programming Languages, Compiler Construction, Data Mining, Machine Learning.

RESEARCH EXPERIENCE

Distributed Computing Systems Group, UMN, Twin Cities **Nov 2017 - Present**
Project: *Geo-distributed Real-Time Data Analytics*
Mentors: *Abhishek Chandra, Jon Weissman*

- Design of algorithms for optimizing Timeliness, Accuracy and Latency in Geo-distributed Real-Time data analytics.
- Currently exploring Apache Flink for implementation of these algorithms.

ADAPT Lab, BITS-Pilani **Apr 2013 - Oct 2014**
Project: *A New Distributed Computing Framework for Data Mining*
Mentors: *Navneet Goyal, Poonam Goyal, Sundar Balasubramaniam*

- Designed and implemented data mining algorithms such as OPTICS, SLINK, DBSCAN for shared memory and distributed memory models.
- Used data distribution and task parallelism techniques for exploiting multicore and multinode architectures. Implemented using OpenMP and OpenMPI libraries in C.
- The work resulted in a number of publications. (See below)

PUBLICATIONS

P Goyal, JS Challa, **D Kumar**, N Goyal, Sundar B. *Grid-R-tree: A data structure for efficient neighborhood and nearest neighbor queries in data mining*, submitted for review in Elsevier DKE.

D Kumar, P Goyal, N Goyal. *An Efficient method for Batch Updates in OPTICS Cluster Ordering*, to appear in International Journal of Data Analysis Techniques and Strategies.[Link]

P Goyal, S Kumari, A Sood, **D Kumar**, Sundar B, and N Goyal. *Exact, Fast and Scalable Parallel DBSCAN for Commodity Platforms*, in International Conference on Distributed Computing and Networking (ICDCN), 2017.[Link]

P Goyal, S Kumari, S Sharma, **D Kumar**, V Kishore, Sundar B, and N Goyal. *A fast, Scalable SLINK Algorithm for Commodity Cluster Computing Exploiting Spatial Locality*, in IEEE International Conference on High Performance Computing and Communications (HPCC), 2016.[Link]

P Goyal, S Kumari, **D Kumar**, Sundar B, N Goyal, S Islam, and JS Challa. *Parallelizing OPTICS for Commodity Clusters* in International Conference on Distributed Computing and Networking (ICDCN), 2015.[Link]

P Goyal, S Kumari, **D Kumar**, Sundar B, and N Goyal. *Parallelizing OPTICS for multicore systems* in ACM India Computing Conference (ACM COMPUTE), 2014.[Link]

Professional Experience

Several Startups **Apr 2016 - Aug 2017**
Technology and Strategy

- Designed and implemented the entire back-end for three startups from scratch.
- The entire back-end functionality was exposed using RESTful APIs implemented using Django web framework and hosted using Amazon web services.
- Gained valuable experience in building scalable and secure back-ends for web and mobile applications.

Goldman Sachs, Bengaluru, India **Nov 2014 - Apr 2016**
Software Developer, Investment Management Division

- Improved the efficiency of risk-management system by suggesting improvements to the SQL queries going to Sybase IQ database.
- Assisted in migrating from Sybase IQ database to MemSQL database for faster access.
- Wrote APIs for accessing MemSQL database.
- Implemented a H2-database based server for allowing real-time updates to the tables residing in the servers.
- Learnt about the real life use-cases of databases.

CSIR-CEERI, Pilani, India **May 2012 - July 2012**
Machine Learning Intern

- Studied, compared and implemented various unsupervised machine learning algorithms.
- Learnt about the use of these algorithms in real world applications

Selected Academic Projects

Win-Loss Prediction for Chess using Machine Learning **Nov 2017 - Dec 2017**

- Training data comprised of endgames with 6 or lesser number of pieces remaining. For such endgames, Nalimov Tablebases have win-loss score for every possible configuration.
- Trained a variety of models including logistic regression, Neural Networks.
- Achieved an accuracy of 85% on the validation dataset.

Restaurant Recommender System **Oct 2013 - Nov 2013**

- An application which can recommend suitable restaurants based on user inputs of location, type of cuisine, type of meal, etc.
- Restaurant reviews taken from yelp.com and processed using NLP techniques.

Complier Construction for a Toy Language **Jan 2013 - Apr 2013**

- Designed lexical, syntax, semantic, code generation phases of compiler in C.
- Efficient use of Hash Tables for constructing symbol tables.

Technical Skills

- **Programming:** C, Java, Python, OpenMPI, OpenMP, MySQL, Verilog, Matlab

- **Mobile and Web Technologies:** HTML, CSS, JavaScript, AngularJS, Django, Android
- **Cloud platforms:** Amazon web services

Honors and Awards

- Awarded a four-year 3M Science and Technology Fellowship for pursuing PhD at the University of Minnesota, Twin Cities. Fellowship covers full tution fees and annual stipend.
- Awarded merit scholarship of total worth Rs 4,75,000 for being in top 10 students among 800 students of BITS, Pilani by the institute. [**Aug, 2010 - May 2014**]
- Awarded research incentive fellowship of Rs 25,000 in recognition of the contribution in the undergraduate thesis project. [**May 2014**]